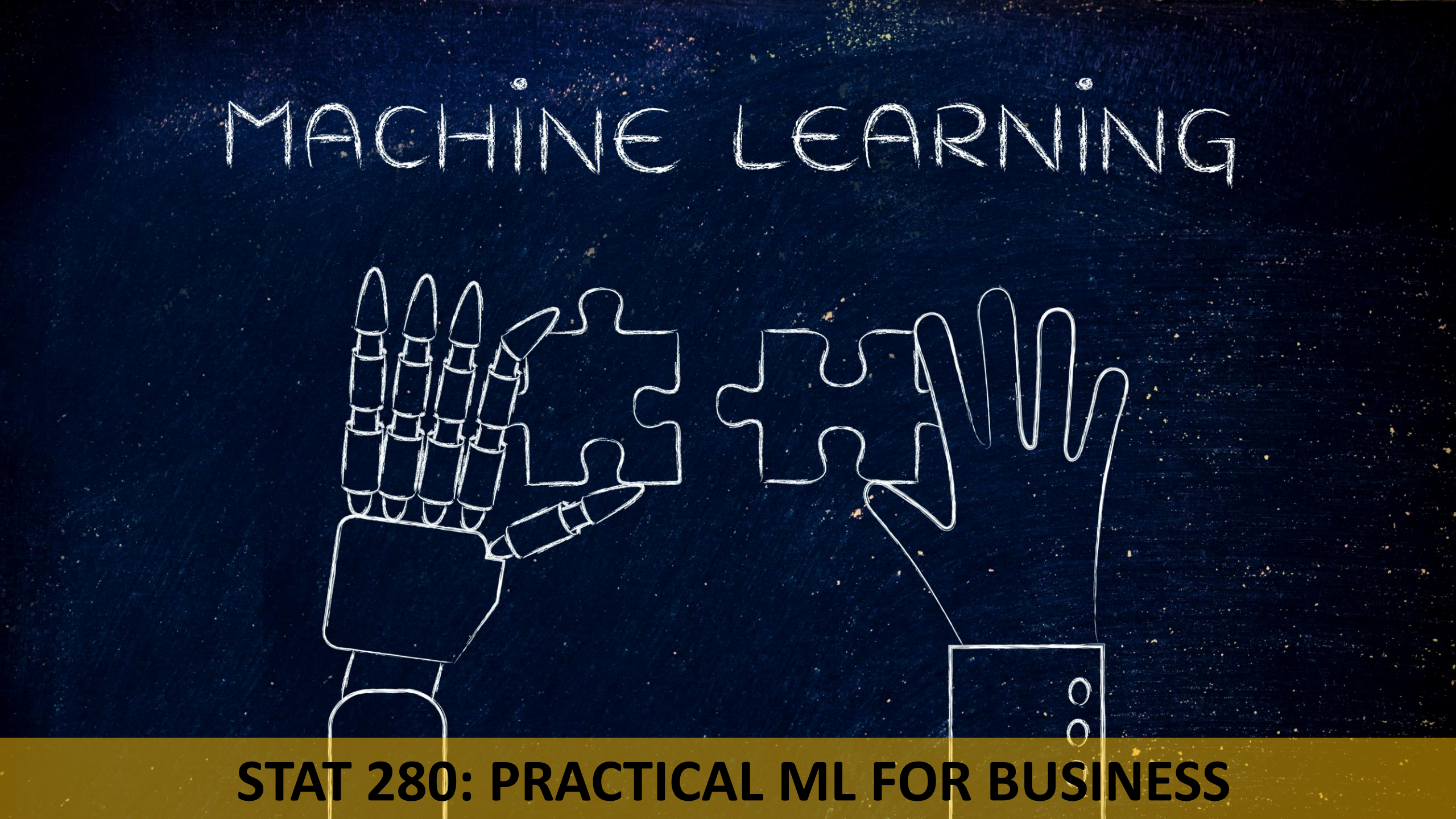


STAT 280: PRACTICAL ML FOR BUSINESS

ACADEMIC INTEGRITY

As a student of the University of the Philippines, I pledge to act ethically and uphold the value of honor and excellence.

I understand that suspected misconduct on given assignments/examinations will be reported to the appropriate office and if established, will result in disciplinary action in accordance with University rules, policies and procedures. I may work with others only to the extent allowed by the Instructor.

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

COPYRIGHT NOTICE

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines".

The University does not authorize you to reproduce or communicate this material. The Material may contain works that are subject to copyright protection under RA 8293. Any reproduction and/or communication of the material by you may be subject to copyright infringement and the copyright owners have the right to take legal action against such infringement.

Do not remove this notice.

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Week 1: Introductions

- Class Introductions and expectations setting
- Course outline and project overview
- Anaconda installation overview

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

About the Instructor

Joby Batiller

- Data Scientist, Gcash (2019)
- > 5 years in Machine Learning
  - Primarily focused on Risk/Fraud
  - Identity resolution
- 5 years instructor in ADMU
- MS Eng’g Math from UoB
- BS PSMSE from ADMU

Introduce Yourself

- Name
- Current job & company
- What new hobby/life realization did you pick up ~~during~~ after the pandemic?

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Stat 280 Course Outline

Week	Date	Topics
1	Jan 20, 2025	0. Introduction • Course outline and project overview • Anaconda installation, GitHub, Slack overview
2	Jan 27, 2025	I. Overview • Introduction to AI/ML • Data Science Lifecycle • Use Case Introduction
3	Feb 3, 2025	II. Jupyter Notebook and Python Basics • Jupyter Notebook introduction • Python 101
4	Feb 10, 2025	III. Use Case 1: Fraud Detection (Supervised) • Understanding Fraud Types • Fraud Detection Techniques • Modeling & Evaluation • Scoring & Deployment & Monitoring

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Stat 280 Course Outline

Week	Date	Topics
5	Feb 17, 2025	IV. Use Case 1: Fraud Detection (Unsupervised) • Data Exploration & Pre-processing • Anomaly Detection • Clustering • Social Network Analysis (SNA)
6	Feb 24, 2025	V. Use Case 1: Fraud Detection (Hands-on) • Data Exploration & Pre-processing • Random Forest • Predictive modeling & Evaluation
7	Mar 3, 2025	VI. Use Case 1: Fraud Detection (Hands-on) • Data Exploration & Pre-processing • Anomaly Detection • Predictive modeling & Evaluation

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Stat 280 Course Outline

Week	Date	Topics
8	Mar 10, 2025	VII. Use Case 2: Recommender Systems • Types of Recommender Systems • Evaluating Recommender Systems • Deployment • Case Studies
9	Mar 17, 2025	VIII. Use Case 2: Recommender Systems (Hands-on) • Data pre-processing • Modeling & Evaluation • Generating Predictions
10	Mar 24, 2025	IX. Use Case 3: Conversational Chatbot • Bot Introduction • Major Use Cases • Bot Building Overview • NLP & NLU Overview • Use Case Discussion

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Stat 280 Course Outline

Week	Date	Topics
11	Apr 7, 2025	X. Use Case 3: Conversational Chatbot (Hands-on) • Rasa Installation • Concepts Review • Rasa Introduction • Hands-on
12	Apr 21, 2025	XI. Use Case 3: Conversational Chatbot (Hands-on cont) • Production Considerations • Connecting to Slack • Chatbot Analytics • Other Topics
13	Apr 28, 2025	Hackathon Day 1
14	May 5, 2025	Hackathon Day 2
15	May 12, 2025	Demo Day
16	May 17, 2025	Demo Day (Delayed)

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Hackathon & Demo Day

- Final project is 30% of final grade
- Work in trios (each with assigned roles)
- For the hackathon, you will have access to various datasets wherein you will pick one, and applying 1 or some of the ML algos we have studied, and come up with an industry specific use case
- Demo day – 15 min presentation and 5 min Q& A
- Panel of Industry Judges
- Winner of Demo Day will get 100% for Group Project Score

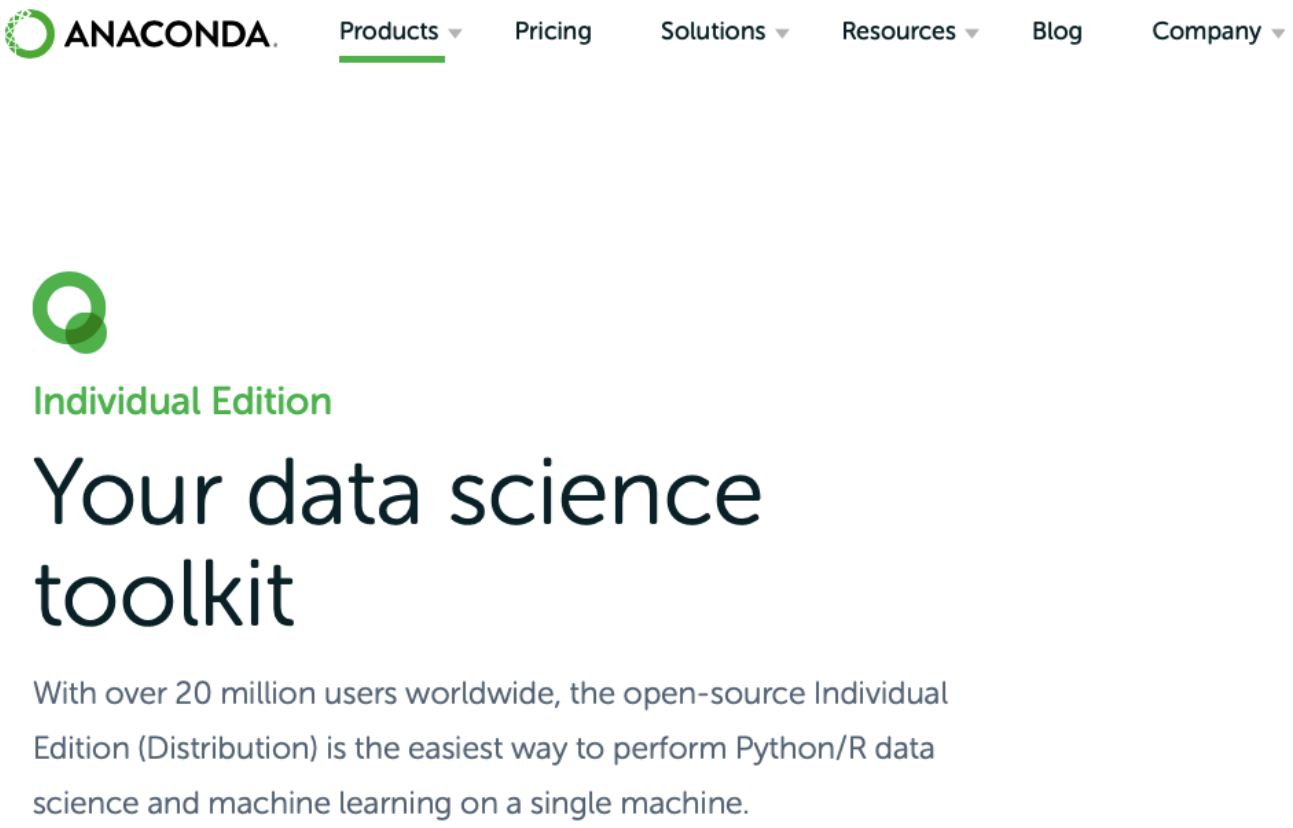
This material has been reproduced and communicated to you by or on behalf of University of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Class Policies

- No exams! ☺ ☺ ☺
- Required forum discussions (4 *discussions – post and 2 comments, 25 points each)*
- Required technical graded assessments (*1 for each use case*)
- Participation in class discussions
  - Camera open at beginning of the class
  - Class picture end of class for attendance
  - It is expected that you will read the LO beforehand to prepare for discussions during synchronous sessions.

of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

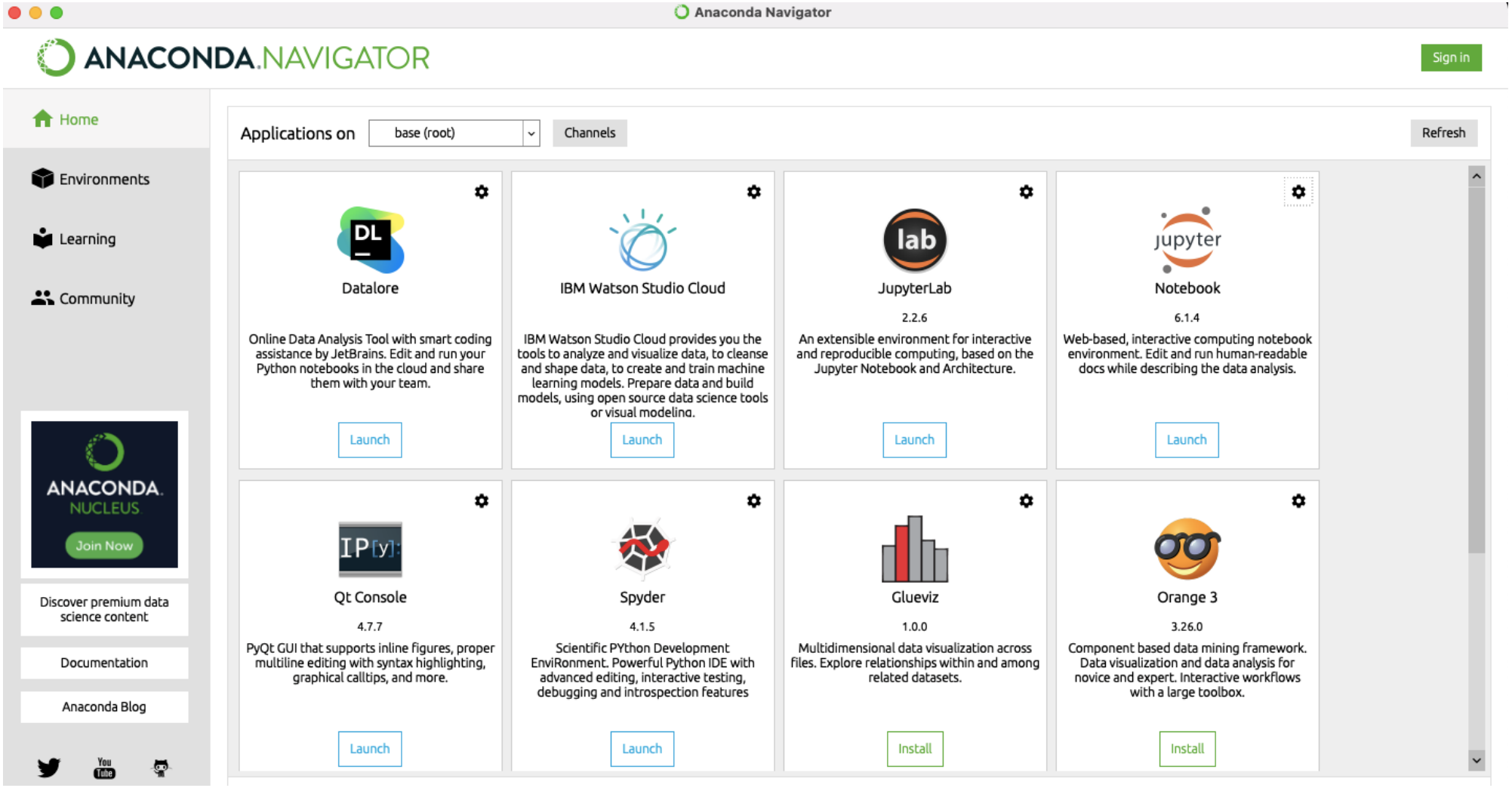
Anaconda Installation Overview



- https://www.anaconda.com/products/individual
- https://docs.anaconda.com/anaconda/user-guide/?utm_source=anaconda.com&utm_medium=individual-get-started
- https://anaconda.cloud/tutorials/8d29a356-46f8-4c5f-9fe8-3b3458b5a252%3Fsource%3Dindividual_tutorial

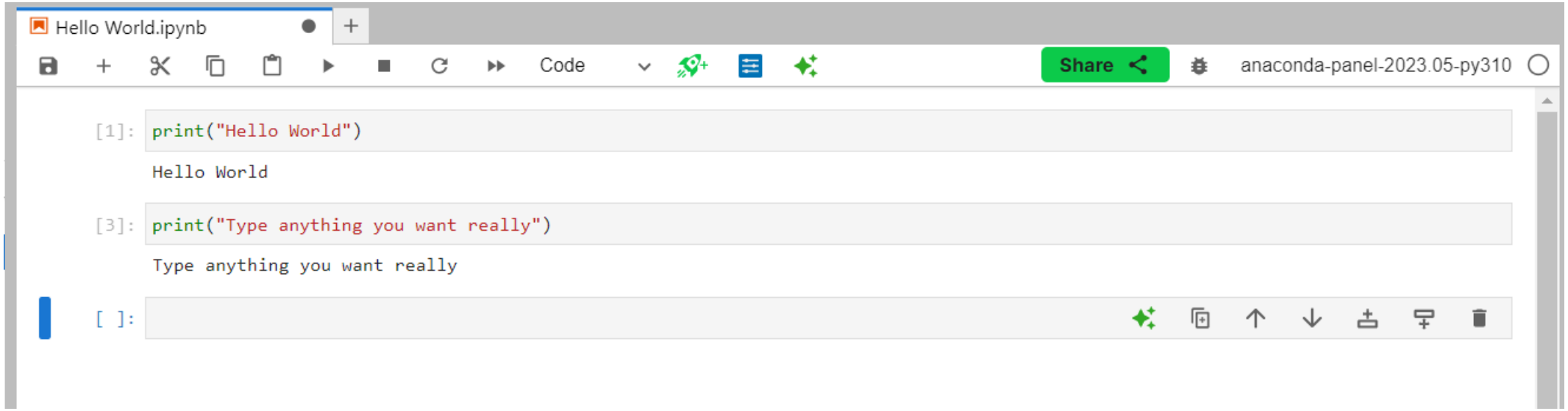
This material has been reproduced and communicated to you by or on behalf of University of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Anaconda Navigator



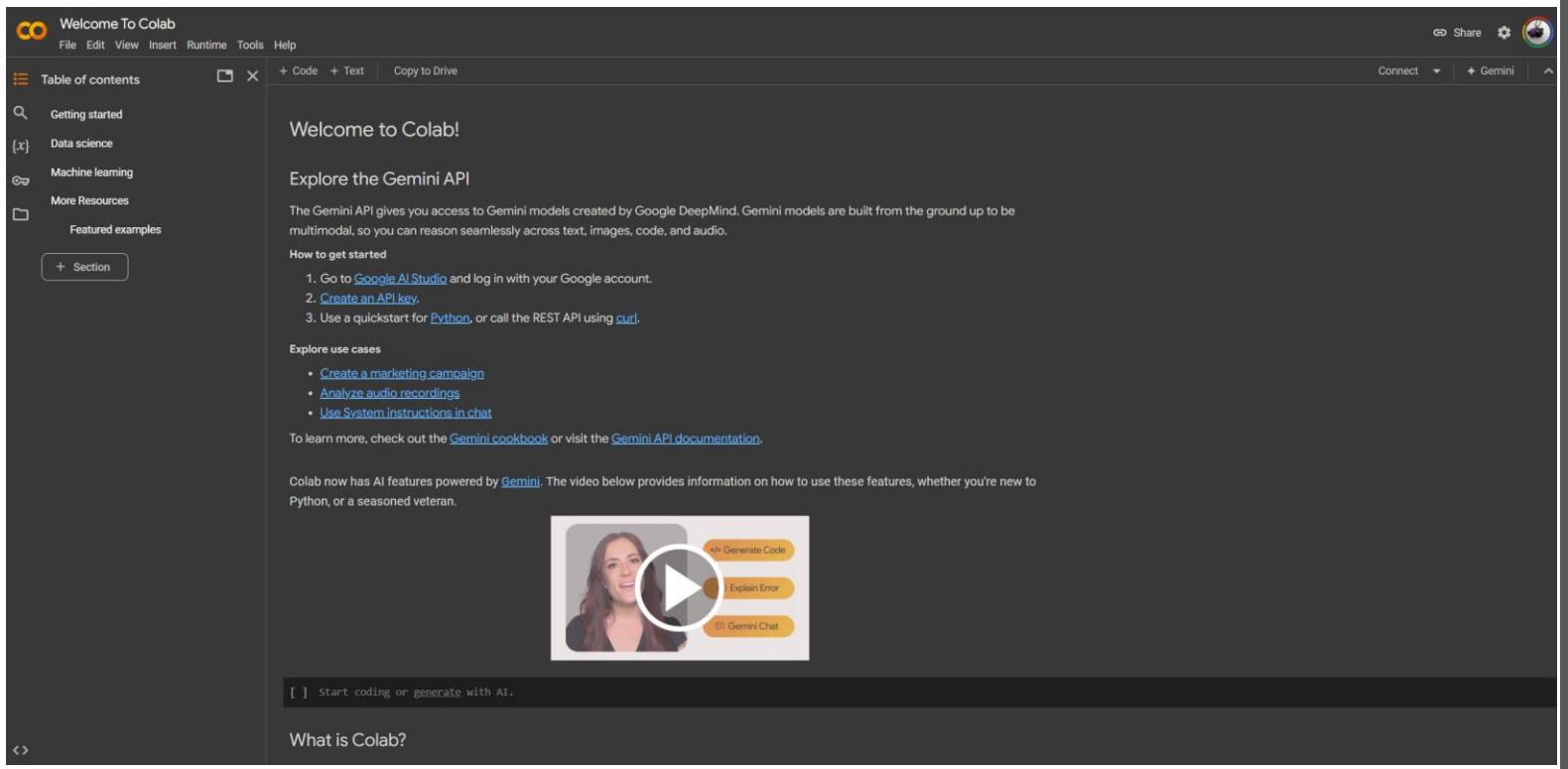
This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Hello World!



This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Google Collab



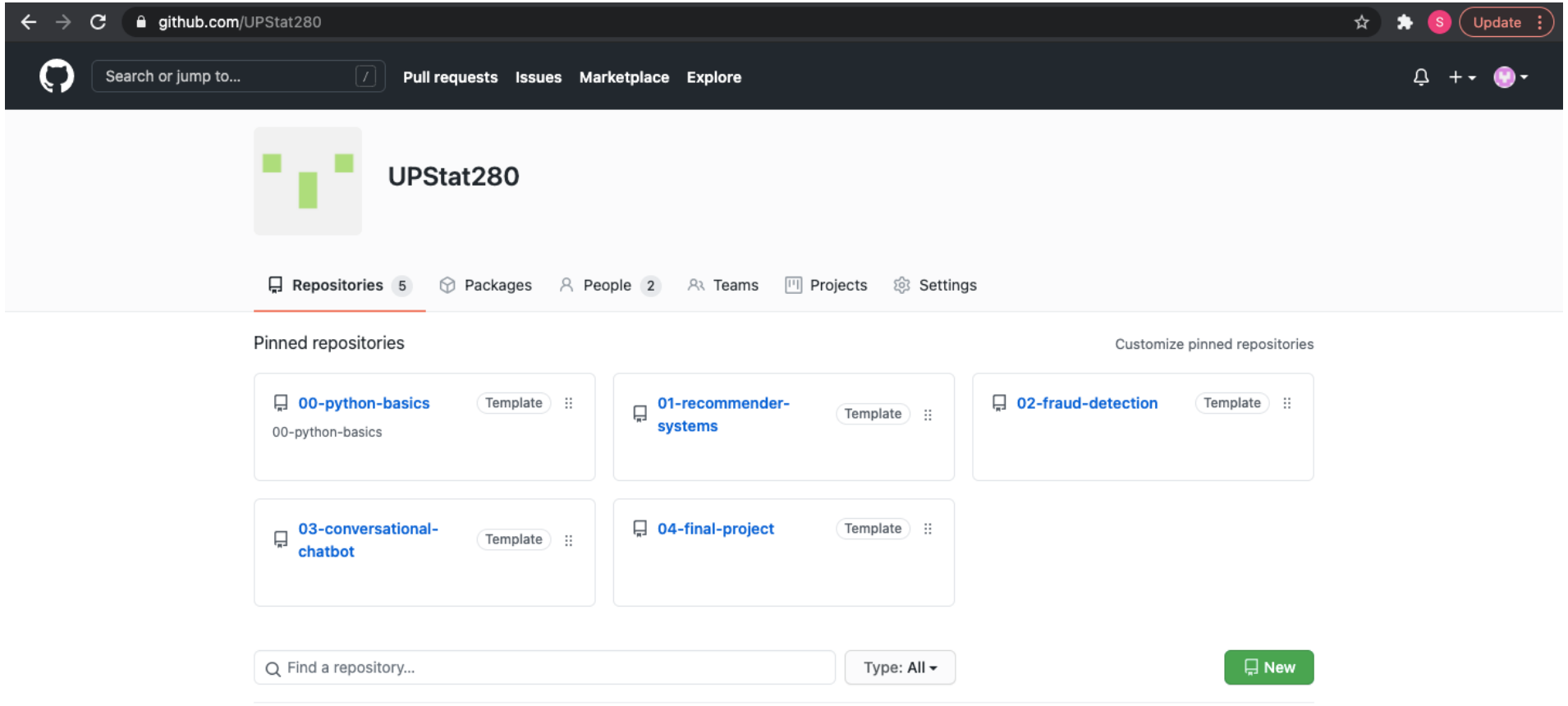
- https://colab.research.google.com/
- Cloud Based Jupyter Notebook Environment
- No setup
- Can link to GDrive

This material has been reproduced and communicated to you by or on behalf of University of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

GitHub Code repository

https://github.com/UPStat280

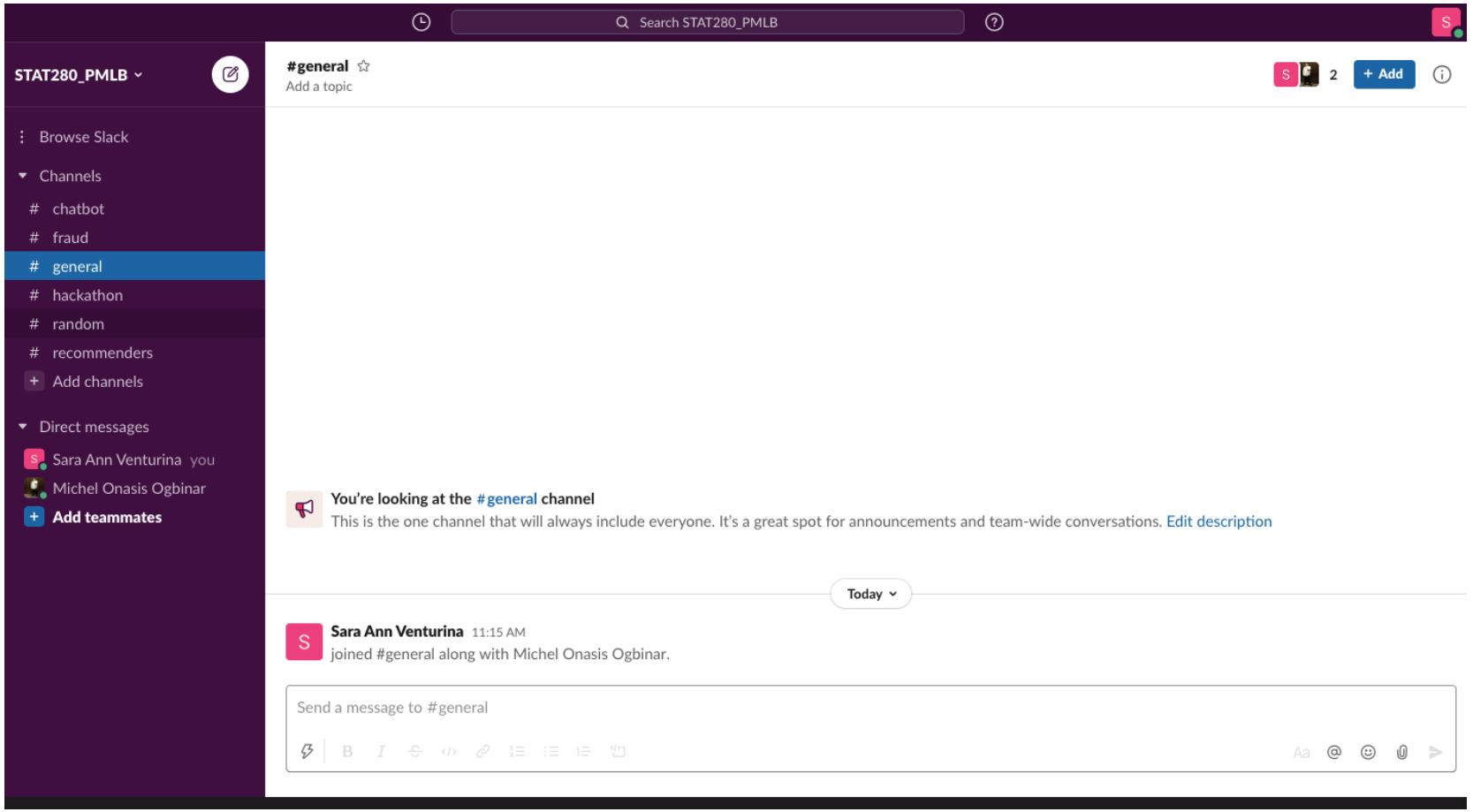
GitHub is a code hosting platform for version control and collaboration.



This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Slack

Slack is a channel-based messaging platform.



👋 Let's move this to Slack! It's a faster, simpler way to talk shop, share files, and work together. You can sign up here: https://join.slack.com/t/stat280mzab2n-wyb6504/shared_invite/zt-2xy7b9hs7-lOKRMkifpvpeBow1T0DRWQ

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Assignment

- Install Anaconda platform on your machine
- Join our Slack channel

of
Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Graded Assessment

- Submit a screenshot of your Hello World program in Slack (unless you can post images on UVLe tell me how)
- Deadline is Feb 3, 2025

of
Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Thank you!

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Stat 280 Main References

- Grus, Joel. Data Science from Scratch: First Principles with Python. O'Reilly Media,2015
- Kane, Frank. Building Recommender Systems with Machine Learning and AI. Sundog Education, 2018
- Baesens, Bart, Veronique Van Vlasselaer and Wouter Verbeke. Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques. Wiley & Sons, 2015
- Shevat, Amir. Designing Bots. Oreilly Media Inc, 2017

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.

Stat 280 Additional References

- Falk, Kim. Practical Recommender Systems. Manning Publications Co, 2019
- Freed, Andrew. Conversational AI. Manning Publications Co, 2021
- Sterne, Jim. Artifical Intelligence for Marketing.J ohn Wiley and Sons, Inc, 2017

This material has been reproduced and communicated to you by or on behalf of University of the Philippines pursuant to PART IV: The Law on Copyright of Republic Act (RA) 8293 or the "Intellectual Property Code of the Philippines". The University does not authorize you to reproduce or communicate this material.